

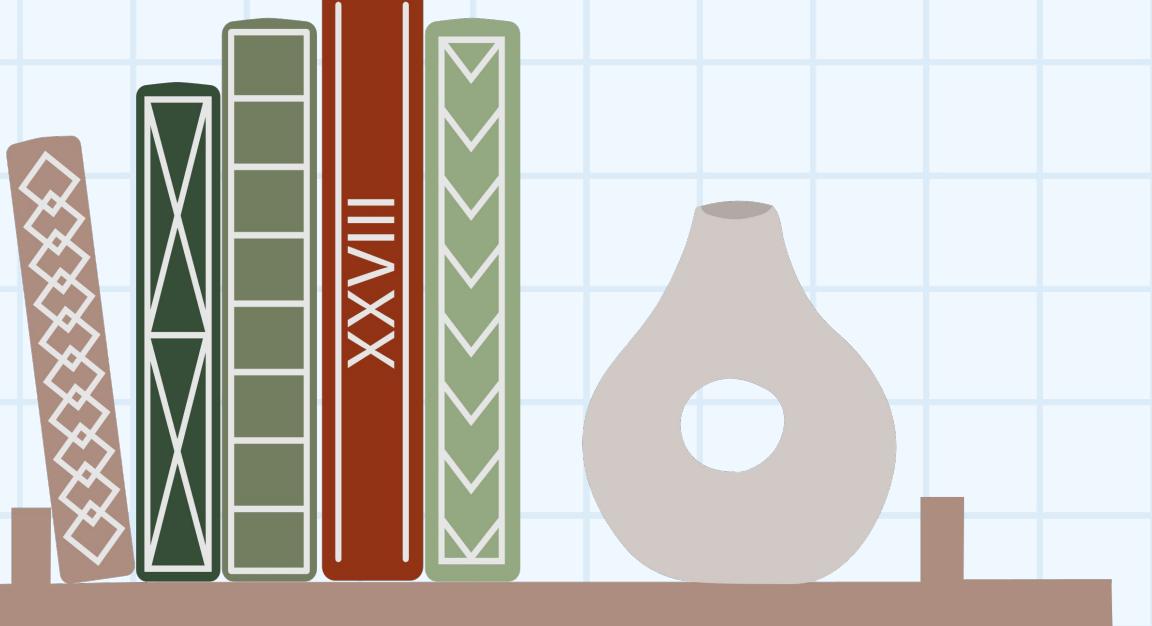
Powered by  
**Futurense**



# BS./BSC.IN

## Applied AI and Data Science

# Basics of Data Analytics



# Let's dive into and learn:



Powered by  
  
**Futurense**

Types of Data Storage

1

Databases

2

Relational Databases

3

# Data Repository



Powered by  
  
**Futurense**

- Data Repository is a broad, general term
- Refers to data collected, organized, categorized and stored for processing and analysis
- Long term preservation and storage of data is an essential part of data analytics
- Stored in a way that is easy to access when required



# Data Repository



Powered by  
  
Futurense

- Having a data repository can help organizations streamline their business decision-making
- Provides a consolidated space to store data critical to your operations
- Easier data-access and troubleshooting
- Efficient reporting and analytics





Powered by



Futurense

# Data Repository

- Suppose you work for a retail organization and have multiple storefronts
- Want to know which store front incurs the highest cost
- Need information on the rents on each unit, utility bills , security costs. Exclude employee salaries or other business expenses.
- Having this data in one place will enable you to quickly find the answer.



# Data Repository



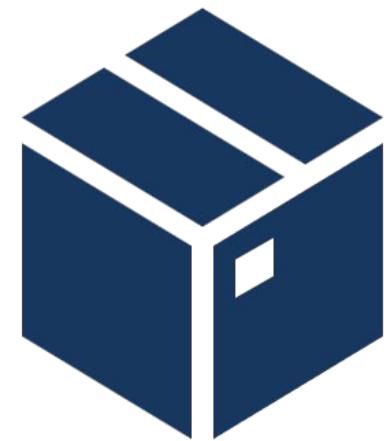
- Data Repository includes
  1. Databases
  2. Data warehouses
  3. Data Lakes
  4. Data Marts



# Databases

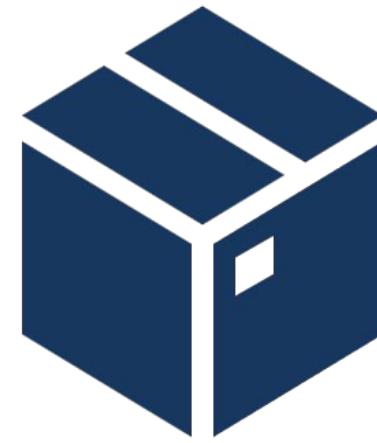


- A data storage system that stores data from various sources for analysis and reporting
- A collection of data created for input, storage, search and retrieval of data
- Stores data in an organized format, typically on a server
- It is used to store data collected from applications and other sources.



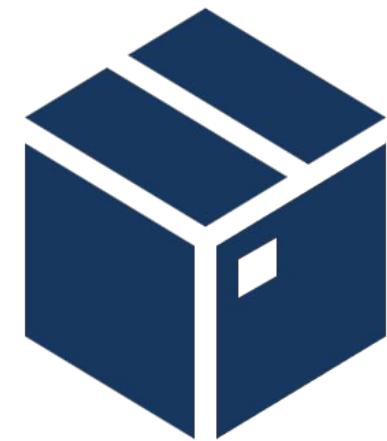
# Databases

- Databases are commonly used for data management tasks such as data validation, data manipulation and data retrieval
- Often used for online transaction processing systems (OLTP), such as customer and sales data for a business



# Databases

- Different Types of databases.
- Choice of database depends on the following factors
  - Data type
  - Data structure
  - Querying mechanism
  - Latency requirements
  - Transaction speeds
  - Final intended use of data



# Relational Databases



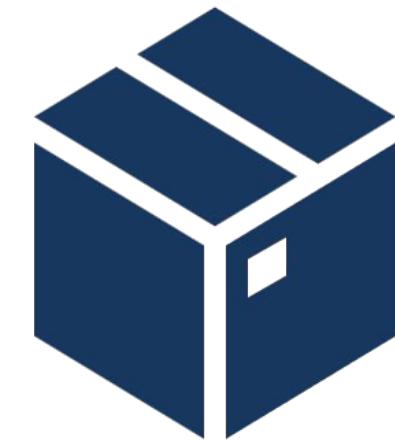
- Data is organized in a tabular form in rows and columns
- Tables can be linked or related based on data common to both
- Typically, data have a well-defined structure
- Optimized for data operations and querying
- Use SQL as the standard querying language



# Relational Databases



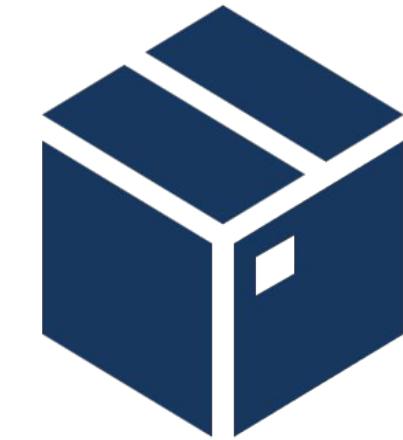
- Data is organized across two or more linked tables
- Each table has a unique set of rows and columns
- Relationships can be defined between tables
- Allows us to retrieve an entirely new table from data in one or more tables with a single query.
- Allows us to understand the relationships among all available data



# Relational Databases



- Minimizes data redundancy
- Allows us to restrict certain fields to certain types and values
- Minimizes data inconsistency and irregularity
- Reduces time required for data cleaning



# Relational Databases



Customer ID	Product ID	Purchase Date
34987	453	01-02-2025
35671	984	12-01-2025
39823	382	29-01-2025
36005	226	05-01-2025

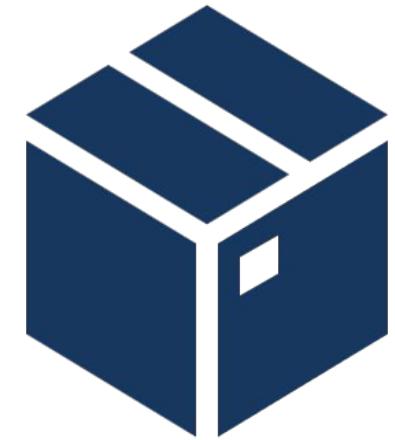
Product ID	Product Name	Price per KG
453	Banana	14
984	Potato	22
382	Watermelon	16
226	Grapes	29

Customer ID	First Name	Last Name	Payment Mode
34987	Rahul	Bajaj	UPI
35671	Neha	Singh	Cash
39823	Uday	Chaudhury	Card
36005	Priya	Agarwal	UPI



# Relational Databases

- Built on the organizational principles of structured data in flat files such as spreadsheets
- But spreadsheets have limited rows and columns
- Relational databases can hold large volumes of structured data
- Ideal for the optimized storage, retrieval, and processing of data for large volumes of data

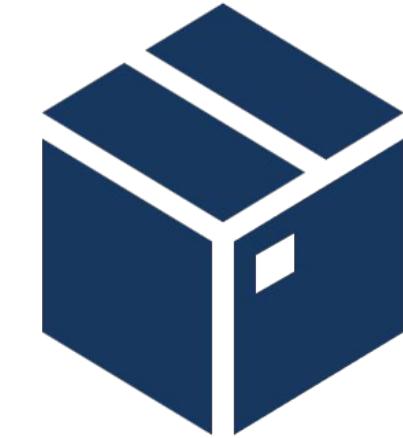


# Advantages of Relational



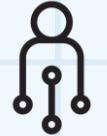
Powered by  
  
Futurense

- Getting insights from joining tables
- Flexibility to make changes while database is in use
- Minimize data redundancy
- Ease of data protection and data recovery





Powered by



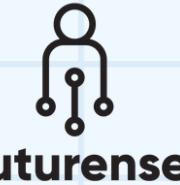
Futurense

# Examples of Relational

- **Online Transaction Processing Applications (OLTP)** - type of database system that manages and processes business transactions in real-time
  - Transaction oriented tasks that support large number of users
  - Run at high rates
  - Manage small amounts of data
  - E.g. banking transactions, e-commerce



Powered by

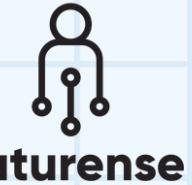


# Disadvantages of Relational

- Does not work well with semi-structured or unstructured data
- Migration between two relational databases is only possible if they have the same structure or schemas, and types
- Some fields have pre-specified value formats and length.  
Can't input more than that.



Powered by



# Non-Relational Databases

- Data can be stored in an unstructured, schema-less manner
- Created to handle the wide variety, diversity, volume and speed of data we see today
- Built for speed, flexibility and scale
- Popularly used to process Big Data

# Database Management Systems



Powered by  
  
Futurense

- Database Management Systems (DBMS) refers to a set of programs that are used for creating and maintaining the database, storing, modifying, and extracting information from the database.
- Done through the process of **querying**

# Recap



Powered by  
  
**Futurense**

## Types of Data storage

### Relational Databases



Powered by



Futurense

# Thank you

